Citation: Clinical and Translational Gastroenterology (2014) 5, e48; doi:10.1038/ctg.2013.20

www.nature.com/ctg

Development and Validation of a Symptom Scale Specific for Ascites Accompanied with Cirrhosis: The ASI-7

Yoshihiro Onishi, PhD, MPH[1,2], Takafumi Wakita, PhD[1,3], Shinichi Fukuhara, MD, FACP, DMSc[2], Yoshinori Noguchi, MD, PhD[4], Mitsuru Okada, MS[5], Isao Sakaida, MD, PhD[6], Yutaka Sasaki, MD, PhD[7] and Kenji Kobayashi, MD[8]

OBJECTIVES: Ascites markedly affects the quality of life of patients with cirrhosis; however, there is currently no scale to measure the symptoms of ascites. We developed a scale to measure ascites-specific symptoms according to psychometric procedures.

METHODS: A team consisting of specialists developed constructs representing the symptoms of ascites and question item pool. The constructs were verified in a qualitative study involving a small number of patients. The item pool was improved through a pilot study, and a prototype of the scale was prepared. To establish the scale and assess its properties, a questionnaire survey was conducted on 175 patients with ascites accompanied with cirrhosis.

RESULTS: On the basis of the results of factor analysis and item response theory–based analyses, seven items, covering a wide range of severities and diverse symptoms, were selected to comprise the final scale (Ascites Symptom Inventory-7; ASI-7). The ASI-7 had a unidimensional factorial structure and high reliability (Cronbach's α coefficient of 0.96). The scale score was correlated with the degree of ascites evaluated by physicians, Short Form-36 (SF-36) physical functioning (PF), and SF-36 vitality (VT; $P<0.001$ each), indicating the criterion validity. The responsiveness after treatment was demonstrated by the mean standardized response of 1.18. Moreover, responses in the scale score were correlated with those in the degree of ascites, body weight, SF-36 PF, and SF-36 VT, respectively ($P<0.001$ each).

CONCLUSIONS: An ascites-specific symptom scale was developed and its reliability, validity, and responsiveness were demonstrated. This simple scale may be used for the evaluation of ascites treatment and monitoring of treatment responses in patients with ascites.

Clinical and Translational Gastroenterology (2014) **5,** e48; doi:10.1038/ctg.2013.20; published online 6 February 2014
Subject Category: Liver

INTRODUCTION

Ascites is one of the primary complications of cirrhosis.[1] During the course of cirrhosis, ascites develops in 30–50% of patients within 5–10 years.[2] As it markedly affects the quality of life (QOL) of patients with cirrhosis,[3,4] treatments relieving ascites may greatly improve patients' feelings.

Diuretics is used for ascites control and paracentesis is performed to manage refractory ascites,[5,6] while new therapeutic techniques, such as albumin infusion and vasopressin receptor antagonists, are being developed.[7–10] In the development of treatments for ascites, objective measurements, such as body weight, abdominal circumference, and ascites volume measured on ultrasonography, have been used to evaluate patients' responses. However, as the primary goal of ascites treatment is to relieve its symptoms, it is worthwhile to add a measurement of ascites-related symptoms for the assessment of ascites treatment.

Feelings regarding the health of patients with cirrhosis have been evaluated using generic QOL scales, such as the Medical Outcomes Study Short Form-36 (SF-36)[11] and sickness impact profile.[12] There are limitations to specifically evaluate the ascites treatment responses using these global scales. The Hepatitis Quality of Life Questionnaire,[13] Chronic Liver Disease Questionnaire (CLDQ),[14] Liver Disease Quality of Life Questionnaire,[15] and Liver Disease Symptom Index 2.0 (LDSI 2.0)[16] have been developed as disease-specific QOL scales focusing on chronic liver disease. These have been used to estimate the disease burden of hepatitis C to assess interferon treatment. Of these disease-specific scales, the CLDQ and LDSI 2.0 involve question items regarding abdominal symptoms; however, the aim of these questions is not to assess the symptoms of ascites. Thus, there is currently no tool that measures the symptoms of ascites despite its marked impact on patients.

We developed an ascites-specific symptom scale through a careful qualitative study and rigorous psychometric procedures, and assessed the reliability and validity, as well as the responsiveness, of this scale.

[1]Institute for Health Outcomes and Process Evaluation Research (iHope International), Kyoto, Japan; [2]Department of Healthcare Epidemiology, Kyoto University, Graduate School of Medicine and Public Health, Kyoto, Japan; [3]Department of Sociology, Kansai University, Osaka, Japan; [4]Department of General Medicine, Japanese Red Cross Nagoya Daini Hospital, Nagoya, Japan; [5]Department of Clinical Research and Development, Otsuka Pharmaceutical, Osaka, Japan; [6]Division of Gastroenterology, Yamaguchi University, School of Medicine, Yamaguchi, Japan; [7]Department of Gastroenterology and Hepatology, Kumamoto University School of Medicine, Kumamoto, Japan and [8]St Luke's Medilocus, Tokyo, Japan
Correspondence: Yoshihiro Onishi, PhD, MPH, Institute for Health Outcomes and Process Evaluation Research (iHope International), Akinono-cho 513, Nakagyo-ku, (Kyoto Daiichiseimei Izumiya Bldg, 7F), Kyoto City, Kyoto 604-0847, Japan. E-mail: onishi@i-hope.jp
Received 29 October 2013; revised 29 October 2013; accepted 14 November 2013

METHODS

This study consisted of two parts (Figure 1). An item pool was prepared through qualitative process in part I. In part II, a quantitative study was conducted to establish a scale and examine its reliability, validity, and responsiveness. The study protocol conformed to the Declaration of Helsinki and was approved by the Ethics Review Board of Ofuna Chuo Hospital on 5 December 2011.

Part I: development of the item pool for the symptom scale. A development team including gastroenterologists, internists, and psychometric experts was organized. Consulting the results of a previous survey involving gastroenterologists and patients, which had been conducted by the members of this team, the team devised constructs (domains) and a sufficient number of items appropriate for measuring ascites-related symptoms.

A qualitative study was then performed to ascertain the appropriateness of the proposed domains. The subjects were three patients with cirrhosis and ascites (two men, one woman, age: 57–69 years) who had consulted a member of the team. Individual patients were interviewed for 60 min in a semistructured manner regarding their history of ascites, symptoms, activities of daily life, and treatment-related changes in their symptoms. Concepts regarding the symptoms of ascites were extracted from the interview records. On the basis of the findings of this analysis, the validity of the domains was confirmed, and the items were revised or new items were added.

Third, a pilot test was conducted to ascertain the appropriateness of the question items. In addition to the three patients who were interviewed, four with cirrhosis and ascites (four men, age: 51–81 years) newly responded to the scale items. Subsequently, they were asked whether the items were appropriate. On the basis of the results of this pilot test, inappropriate items (questions in which responses converged to particular choices or those that were difficult to understand) were removed, and a prototype of the questionnaire consisting of a sufficient number of questions was prepared for the next phase.

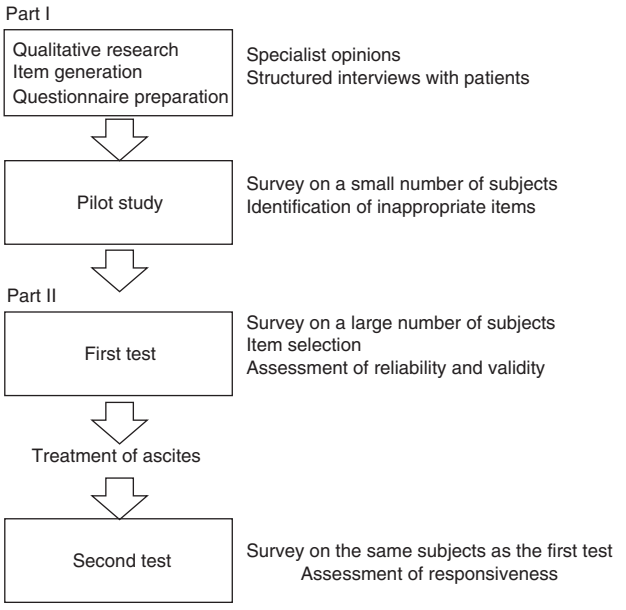

Figure 1 The procedure of scale development in this study. These processes conformed to the standard procedures for the development of a psychological scale.

Part II: establishment of a scale and assessment of its properties

Subjects. A survey involving a number of patients was conducted using the questionnaire developed in part I. Patients with ascites accompanied with cirrhosis in whom the start or switch of treatment for ascites (paracentesis, admission, albumin infusion, or diuretics) was scheduled were eligible. Patients with uncontrollable liver cancer and those with hepatic encephalopathy (grade II or higher) were excluded. A total of 175 patients were registered from 24 institutions between January 2012 and December 2012.

Study procedures. After giving informed consent, patients were requested to respond to the questionnaire including the scale items. Simultaneously, the physical functioning (PF) and vitality (VT) of the Japanese-version SF-36 (ref. 17) were acquired. These have been shown to be closely associated with the presence or absence of ascites among eight subscales of the SF-36.[3] A norm-based score was used as a subscale score of the SF-36. This score was standardized, with the mean in the representative sample of a general population as 50 and s.d. as 10. The attending physicians reported the severity of cirrhosis (Child–Pugh score), degree of ascites, contents of the intended ascites treatment, and body weight. The degree of ascites was evaluated using five grades: none, slight, moderate, marked, and tense. To evaluate the responsiveness of the scale, patients were again requested to respond to the same questionnaire after the completion of the intended treatment (regardless of the treatment response). The attending physicians reported the degree of ascites and body weight.

Questionnaire. The questionnaire consisted of the scale items developed in part I, as well as the SF-36 PF and VT. Patients responded to the ascites scale items using a five-point Likert scale: 0 = not applicable, 1 = neither, 2 = slightly applicable, 3 = applicable to some degree, and 4 = very applicable, according to an instructive sentence, ''Please answer questions with respect to today's symptoms of ascites (bloatedness in the belly) and symptom-related matters''.

Item analysis, factor analysis, and item selection. The mean, s.d., and frequency of missing responses were described for each item. Factor analysis with principal factor method was used to examine the factorial structure among the items. The number of latent factors was examined by the attenuation of eigenvalues. Factor loading, which represented the correlation between the estimated factor and each item, was assessed. After confirming the unidimensionality (one-dimensional factorial structure) of the scale and high values of factor loading, we decided to select several items to establish a final scale. The item response theory (IRT), which has commonly been used to select questions for an achievement test,[18] was applied to evaluate the performance of each item. The Generalized Partial Credit Model[19] was

used as an IRT model. This model estimated three parameters: location, slope, and category. The location parameter represented the severity of symptoms to which an item responded (difficulty). The slope parameter represented the ability of an item to discriminate between mild- and severe-symptom patients (discriminating ability). The category parameter represented the degree of difficulty inherent to respective options of the Likert scale, and was estimated as values common among all items. On the basis of the findings of IRT analysis as well as the meanings of each item, items were selected by the development team's consensus, considering the following points: (1) items responding to mild-to-severe symptoms were evenly included; (2) the slope parameter of the individual items was high; (3) the meanings of the individual items were relevant and specific; and (4) a combination of items covered diverse symptoms (no inclination toward particular domains). The Test Information Function was calculated to examine the precision of the measurement of the final scale.[18]

Reliability and validity. The score of the final scale was calculated by summing up the score of each item (0–4 points), ranging from 0 to 28 points. The following analysis was conducted using the data obtained before the treatment: the distribution of the scale score was described; the reliability (internal consistency) was assessed using Cronbach's α coefficient; criterion validity was assessed by examining the association of the scale score with three external standards: the degree of ascites evaluated by physicians, SF-36 PF, and SF-36 VT.

Responsiveness. To evaluate the responsiveness of the final scale, we investigated the features of changes in the scale score after treatment, as described below. First, to assess the magnitude of the changes, the standardized response mean (SRM)[20] was calculated by dividing the mean of changes in the scale score by the s.d. of the changes. Empirical criteria of large (>0.8), moderate (0.5–0.8), or small (0.2–0.5) were used for interpreting SRM values.[20–22] Second, we investigated the association between changes in the scale score and those in the following external standards: the degree of ascites, body weight, SF-36 PF, and SF-36 VT. To facilitate the interpretation of changes in the scale score, single linear regression analysis was conducted using changes in the scale score as an independent variable and changes in body weight or the SF-36 PF score as a dependent variable. Third, to evaluate the ability of the scale to reflect changes in the external standards the area under the receiver operating characteristic (ROC) curve[23] was calculated for changes in the scale score with those in the degree of ascites and those in body weight as gold standards, respectively. An area under the ROC curve of 0.5 indicated the absence of a discriminating ability, and that of 1.0 indicated complete discrimination. In addition, similar ROC analysis was performed for changes in the SF-36 PF and VT scores as index variables.

Statistical analysis. The distributions of continuous and categorical variables were summarized as the mean (s.d.) and proportion, respectively. Correlation coefficients and/or the analysis of variance with Tukey's multiple comparison were used to analyze the association between the two variables. A *P* value of 0.05 was regarded as significant. SPSS version 21.0 (IBM, New York, NY, USA) and PARSCALE 4.0 (Scientific Software International, Skokie, IL, USA) were used for the factor analysis and IRT analysis, respectively.

RESULTS

Part I. Six domains were devised by the development team: (1) abdominal bloating, (2) restriction on activities, (3) restriction when lying, (4) restriction on eating, (5) fatigue, and (6) others. A total of 41 questions were generated on these six domains. The qualitative study involving patients with ascites indicated (1) that the six domains covered the ascites-related symptoms that patients felt, (2) that patients expressed the site as "the periphery of the navel", and (3) that, to evaluate ascites-specific symptoms, it was necessary to differentiate them from the symptoms of underlying cirrhosis. On the basis of these findings, the development team revised the items and added new ones, and prepared a revised item pool consisting of 51 questions.

In a pilot study with the 51 items, ceiling or floor effects were observed in some items. Furthermore, two items, "it is hard to walk to the toilet", and "my activities are slower than previous ones", were negatively correlated with the other items. Considering the item wording and correlation among the item meanings, 14 items were removed, resulting in 37 items for part II (Supplementary Table 1 online).

Part II

Subjects. The characteristics of 175 subjects of part II are shown in Table 1. Relatively broad distributions of Child–Pugh scores or degree of ascites were noted. The mean score of norm-based SF-36 PF of 12.8, a value lower than the national mean by >3.5 s.d., suggested that the subjects' activities were strikingly affected. The SF-36 VT score was 33.7, which also showed a marked decrease.

Item and factor analyses. There were no marked ceiling or floor effects in any item based on the item analysis (Supplementary Table 1). Missing responses were rare.

Factor analysis showed the attenuation of eigenvalues of 26.9, 1.6, 1.1, and 0.6 (Supplementary Figure 1 online). The first factor explained 73% of the variance in the data, showing

Table 1 Characteristics of subjects in part II of this study ($n=175$)

Female/male	68/107
Age, mean (s.d.) [range]	67 (10) [36–88]
Body weight, mean (s.d.) [range]	63 (10) [36–90]
Child-Pugh class	
A	21 (12%)
B	88 (50%)
C	66 (38%)
Degree of ascites evaluated by physicians	
Slight	55 (31%)
Moderate	72 (41%)
Marked	32 (18%)
Tense	16 (9%)
SF-36 physical functioning, mean (s.d.)[a]	12.8 (18.6)
SF-36 vitality, mean (s.d.)[a]	33.7 (10.8)

SF-36, Short Form-36.
[a]Norm-based score.

a strong unidimensionality. Factor loading was high (0.611–0.937). There were no marked differences in the mean, s.d., or factor loading among the items (Supplementary Table 1), which suggested that these 37 items mutually involved exchangeable contents for measuring ascites-related symptoms. These findings indicated that it was possible to greatly reduce the number of items, and the development team decided to select several items and arrange a scale.

Item selection. When selecting items, two items of the ''others'' domain, of which factor loading was relatively low, were excluded. The results of IRT analysis involving the remaining 35 items are shown in Figure 2 and Supplementary Table 1. The location parameter, which reflected the severity to which each item responded, ranged from -0.65 to 0.54. Many items had negative values and were suitable for relatively mild symptoms. Regarding the domains, abdominal bloating corresponded to the range of the mildest symptom, followed by restriction when lying, restriction on activities, and restriction on eating. The range of measurements for restriction on activities was large, while that for fatigue was small and overlapped abdominal bloating. These results were consistent with the physicians' clinical experiences. Considering the results of this IRT analysis and the contents of the question items covering various symptoms, the team selected seven items for the final scale.

The seven items comprising the scale are shown in Table 2. This scale was named the Ascites Symptom Inventory-7 (ASI-7). These seven items covered various domains and were arranged in a sequence from mild-to-severe symptoms, facilitating clinical interpretation. The scale score of the ASI-7 showed a symmetric distribution within a range of 0–28 points. The mean value (s.d.) was 15.1 (7.8) points. There were no ceiling or floor effects. The seven-item scale score was highly correlated with the total score of the 35 items before item reduction ($r = 0.98$).

Reliability and validity. Cronbach's α coefficient for the ASI-7 was 0.96. The test information function exceeded 10 at a location parameter ranging from -2.0 to 1.5 (Supplementary

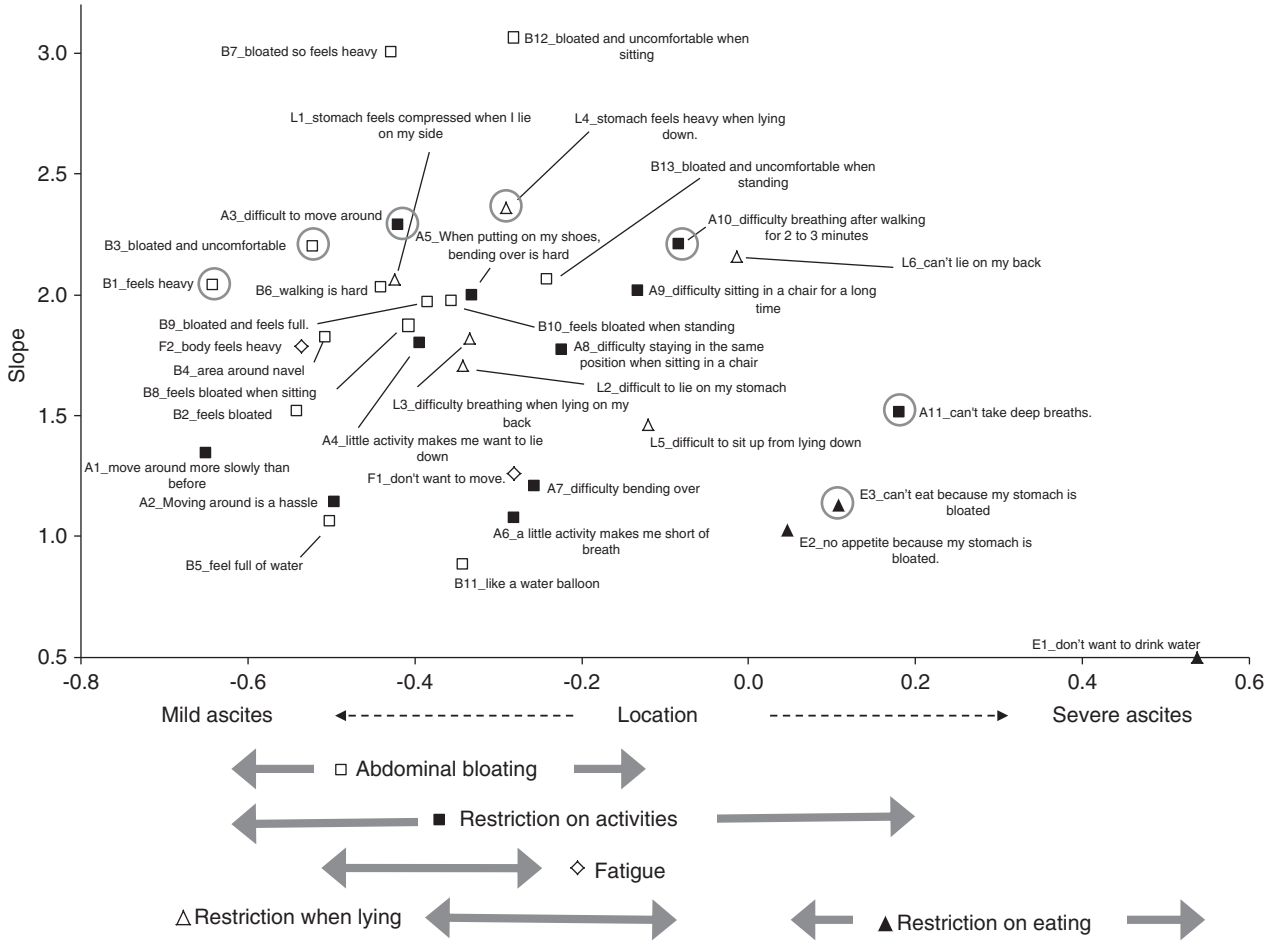

Figure 2 A plot of the properties of each question item estimated by the item response theory (IRT) analysis using the Generalized Partial Credit Model. The horizontal axis represents the location parameter, indicating the disease severity to which items responded. The vertical axis represents the slope parameter, indicating the ability of items to distinguish differences in the disease severity of individuals. Different symbols indicate that items belong to different domains. Refer to Supplementary Table 1 for respective values. The thick arrows below the plot represent the range of location covered by the six domains. respectively. When selecting items to measure various severity of symptoms, the IRT suggests that items should be selected so that they cover a range of locations evenly, and that items with a greater slope should be selected at the same location. The circles in the plot represent seven items selected by the development team for the final scale.

Figure 2). These findings supported the reliability of the scale and confirmed that the precision of the measurement was preserved.

The correlation between the scale score and external standards is shown in Figure 3. The scale score was higher in patients in whom the degree of ascites was more severe. The mean score exceeded 25 points in patients graded to "tense". There were significant differences among the four grades of ascites. The scale score was correlated with the SF-36 PF and VT ($r = -0.63$, $P < 0.001$ and $r = -0.55$, $P < 0.001$, respectively). Analysis of variance and multiple comparisons also confirmed significant correlations (Figure 3).

Responsiveness. In 172 of the 175 patients, data were obtained after treatment. The median interval of data collection between pre-treatment and post-treatment measurements was 28 days (interquartile range: 14–52 days). Of these, paracentesis had been performed in 30 patients (17%), treatment involving albumin infusion in 23 patients (13%), and therapy with diuretics alone in 119 patients (69%). The median intervals of data collection for these treatment groups were 9 days (3–38), 30 days (15–74), and 28 days (14–54), respectively.

The treatment response, evaluated based on changes in the degree of ascites after treatment, were distributed in a two-rank or greater improvement in 38 patients (22%), a one-rank improvement in 93 patients (54%), no changes in 38 patients (22%), and a one-rank deterioration in 3 patients (2%). The mean change (s.d.) in body weight was -2.9 (2.2) kg (range: -13.6 to $+4.0$ kg). The mean (s.d.) SF-36 FP and VT scores increased by 12.2 (14.8) and 8.0 (10.8), respectively, after the treatment.

The mean scale score (s.d.) decreased by 6.7 (5.6) points (range of change: -27 to $+1$) after treatment. The SRM was 1.18, corresponding to a large change (>0.8) according to the empirical criteria. Examining the SRM by different treatment groups (Table 3), the paracentesis group had the highest value. In the diuretic group, the value was 1.06, which still corresponded to a large change.

The correlation between changes in the scale score and those in the external standards is shown in Figure 4. The scale score decreased more (analysis of variance, $P < 0.001$) in groups with more marked improvements in the degree of ascites. Its relationship with changes in body weight was

Table 2 Seven items composing the ASI-7

No	Item	ID	Domain
1	My stomach feels heavy	B1	Abdominal bloating
2	My stomach is bloated and uncomfortable	B3	Abdominal bloating
3	My stomach is bloated, so it is difficult to move around	A3	Restriction on activities
4	My stomach feels heavy when lying down.	L4	Restriction when lying
5	I have difficulty breathing after walking for 2–3 min	A10	Restriction on activities
6	I cannot eat because my stomach is bloated	E2	Restriction on eating
7	I cannot take deep breaths.	A11	Restriction on activities

ASI-7, Ascites Symptom Inventory-7.
Instructions: "Please answer the following questions about your ascites (abdominal fluid or fullness) and related symptoms, as they are today." Likert choices: 0 = does not apply at all; 1 = slightly applies; 2 = somewhat applies; 3 = strongly applies; 4 = very strongly applies.

Table 3 SRM of the ASI-7 by groups undergoing different treatments

Type of treatment	SRM
Paracenteses	1.55
Albumin infusion and diuretics	1.32
Diuretics alone[a]	1.06

ASI-7, Ascites Symptom Inventory-7; SF-36, Short Form-36; SRM, standardized response mean.
[a]Treatment including neither paracenteses nor albumin infusion.

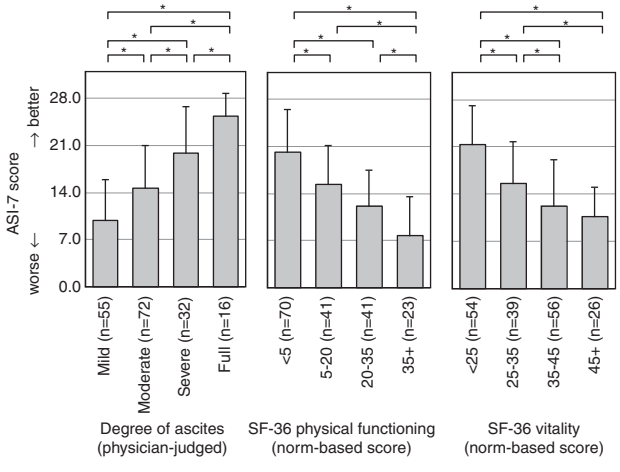

Figure 3 Mean scale score (s.d.) of the Ascites Symptom Inventory-7 (ASI-7) according to the level of external standards. For the Short Form-36 (SF-36) subscales, the subjects were divided into four groups according to arbitrary criteria so that there was no group in which the number of subjects was very small. The ASI-7 score was correlated with the three external standards (analysis of variance: $P < 0.001$ each). Tukey's multiple comparison: $^*P < 0.05$.

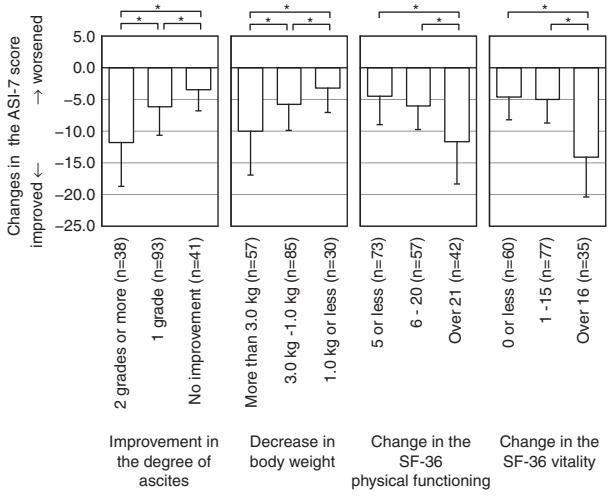

Figure 4 Mean scale score changes (s.d.) of the Ascites Symptom Inventory-7 (ASI-7) according to the level of changes in external standards. The changes in the scale score were correlated with the four external criteria (analysis of variance: $P < 0.001$ each). Tukey's multiple comparison: $^*P < 0.05$.

similar. Changes in the scale score were correlated with those in the SF-36 PF and VT scores ($r = -0.60$ ($P < 0.001$) and $r = -0.64$ ($P < 0.001$), respectively), which were also shown by analysis of variance (Figure 4). Single linear regression analysis revealed that a 2.0 kg decrease in body weight and 10-point improvement in the norm-based SF-36 PF score corresponded to an approximately 5-point decrease in the scale score.

The area under the ROC curve for changes in the scale score with those in the degree of ascites or body weight as gold standards was in a range of 0.7 to 0.8. These values were higher than those for the SF-36 PF or VT scores (Table 4).

DISCUSSION

As a result of the demand for patient-based outcome assessments, the measurement of symptoms, disabilities, and QOL is crucial. Several QOL scales specific to chronic liver disease have been developed.[13–16] However, the symptoms of ascites have not been sufficiently assessed because these QOL scales target a wide spectrum of clinical features. We believe that assessing the symptoms of ascites is essential to evaluate ascites treatment and to monitor the response to the treatment, which promoted us to develop a new scale to measure ascites-specific symptoms. Indeed, in the subject population in part II, PF measured using the SF-36 was remarkably reduced. This indicates that the presence of ascites is a serious burden to patients although factors other than ascites should be combined.

In this study, we developed the ASI-7, a scale to measure symptoms specific to ascites. This simple scale consisted of seven short items covering a wide range of severities and symptoms of ascites, and it had favorable reliability, criterion validity, and responsiveness.

Scale development process. Here, we present several characteristics that are key to developing the ASI-7. First, this scale was developed to quantitatively assess ascites-specific symptoms and its response to treatment using a self-administered questionnaire. Second, before conducting the quantitative study, we postulated several domains to comprehensively cover ascites-related symptoms, and prepared a sufficient number of items to ensure the thoroughness and precision of the scale. These were carefully examined in a qualitative study involving patients to confirm the thoroughness of domains and to assess the understandability and acceptable wording of the items. Thus, we verified the face and content validity of these items, and then proceeded to the next step. Third, careful psychometric analyses were performed on the data collected from patients requiring ascites treatment. Factor analysis indicated strong unidimensionality and high factor loading, which supported the appropriateness of the qualitative study. The substantial correlation among the items allowed us to greatly reduce the number of items. Fourth, data were collected from a number of patients after the treatment, and the responsiveness of the scale was evaluated. The scale showed a high responsiveness, suggesting that the above process was appropriate.

Scale properties. The ASI-7 has some favorable properties. First, sentences were short and understandable, and intuitively interpretable, reflecting face and content validity. Second, high reliability was shown by Cronbach's α coefficient and the test information. Third, the scale scores were distributed throughout a possible range of 0–28 points, suggesting that the measurement range was suited to the severity of the symptoms to be evaluated. Fourth, criterion validity was demonstrated by the scale score correlating well not only with the SF-36 PF and VT scores, which were based on patients' self-evaluations, but also with the degree of ascites evaluated by physicians. Fifth, responsiveness was demonstrated by the large SRM in the study subjects. This was true even in a subgroup treated with diuretics alone. Responses in the scale score agreed with those in the degree of ascites and body weight. These results suggest that this scale may be used for evaluating ascites treatment, including that with diuretics. The ROC analysis confirmed that this condition-specific scale more sensitively detected the responses to ascites treatment than that of the global scales, as expected by the item contents.

The clinically significant response level can be reviewed to some degree, although this study was not designed to clarify it. On single regression analysis, a 5-point improvement in the scale score corresponded to 10 points (i.e., 1.0 s.d.) of the norm-based SF-36 PF score. Some studies involving QOL measurements in patients with chronic liver disease reported that, in patients with decompensated cirrhosis, the SF-36 PF score (0–100 points) was about 15 points lower than that in healthy adults.[16,24] This value corresponded to 0.8 s.d. based on the national norm. Another study compared patients with Child–Pugh class C with those with class A,[3] and indicated that the difference in the SF-36 PF score (expressed by the Z score) between the two groups was approximately one point (i.e., 1.0 s.d.). Thus, an ASI-7 score of five points roughly corresponded to the average reduction in the physical function of patients with decompensated cirrhosis relative to healthy

Table 4 Area under the ROC curve as a measure of the responsiveness of the ASI-7 and SF-36 scales

		Instrument		
External standard	**Definition of a good outcome**	**ASI-7**	**SF-36 physical functioning**	**SF-36 vitality**
Improvement in the degree of ascites	Two grades or more vs. others	0.80	0.66	0.73
	One grade or more vs. others	0.74	0.63	0.66
Decrease in body weight	>3.0 kg vs. others	0.76	0.55	0.65
	>1.0 kg vs. others	0.80	0.70	0.66

ASI-7, Ascites Symptom Inventory-7; ROC, receiver operating characteristic; SF-36, Short Form-36.

subjects or those with compensated cirrhosis; this may represent a significant impact for patients. To assess minimal clinically important differences, further studies are required.

Limitations. In this study, quantitative analysis was conducted in a single population. Therefore, results may have been optimized to the study population. We do not believe that there was a significant bias because the number of institutions was not small ($n = 24$) and patients with various severities of ascites were included. However, validation of the results of study is anticipated in another population. In addition, the generalizability of these items may be limited in countries in which the descriptors of symptoms are different. Cultural adaptation studies may be needed in such cases.

Clinical significance. One may say that body weight measurements are sufficient for the evaluation of responses to ascites treatment. However, because ascites treatment aims to provide relief from the symptoms of ascites, patient-reported outcomes are inevitable for measuring treatment impacts. As this scale consisted of seven short questions, it may be useful as eligibility criteria or the end points in clinical studies and for treatment monitoring in clinical practice. Furthermore, this scale involved questions on diverse symptoms, although the number of questions was small. Therefore, it may universalize and complement patients' words regarding symptoms and function as a physician–patient communication tool.

Conclusions. The seven-item Ascites Symptom Index, a self-administered symptom scale specific to ascites accompanied with cirrhosis, was developed through careful item preparation and rigorous psychometric procedures. The reliability, validity, and responsiveness of the scale were confirmed. This simple scale allows the quantitative assessment of improvements in ascites symptoms by treatment. It may be used for the evaluation of ascites treatment and monitoring of treatment responses in patients with ascites.

CONFLICT OF INTEREST

Guarantor of the article: Yoshihiro Onishi, PhD, MPH.
Specific author contributions: Yoshihiro Onishi planned and conducted the study, collected and interpreted data, drafted the manuscript, and approved the final draft submitted. Takafumi Wakita planned the study, interpreted data, drafted the manuscript, and approved the final draft submitted. Shinichi Fukuhara planned the study, interpreted data, and approved the final draft submitted. Yoshinori Noguchi planned the study, interpreted data, and approved the final draft submitted. Mitsuru Okada planned and conducted the study, collected and interpreted data, and approved the final draft submitted. Isao Sakaida reviewed and approved the manuscript. Yutaka Sasaki reviewed and approved the manuscript. Kenji Kobayashi planned the study, interpreted data, and approved the final draft submitted.
Financial support: Funding for this study was provided by Otsuka Pharmaceutical.
Potential competing interests: None.

Acknowledgments. The following investigators participated in this study: Atsushi Yoshida (Gastroenterology, Ofuna Chuo Hospital); Takayuki Suzuki (Internal Medicine, Ofuna Chuo Hospital); Masamichi Yokoe (General Medicine, Japanese Red Cross Nagoya Daini Hospital); Etsuro Orito, Katsumi Hayashi, Atsunori Kusakabe (Gastroenterology, Japanese Red Cross Nagoya Daini Hospital); Toshifumi Sasanuma (Shimeidouiin Clinic); Fumihiro Iwata (Iwata Naika Iin); Hidehiro Kawabata (Kawabata Medical and Gastrointestinal Clinic); Yasushige Kashima (Surgery, Tochinoki Hospital); Takumi Kawaguchi (Gastroenterology, Kurume University School of Medicine); Takehiro Kurotaki (Akabira Municipal General Hospital); Satohiro Matsumoto (Saitama Medical Center, Jichi Medical University); Makoto Shiraki (Internal Medicine, Gifu University School of Medicine).

Study Highlights

WHAT IS CURRENT KNOWLEDGE

- Ascites markedly affects the QOL of patients with cirrhosis.
- There is no scale to measure the symptoms of ascites.

WHAT IS NEW HERE

- An ascites-specific symptom scale was developed according to psychometric procedures.
- Reliability, validity, and responsiveness were demonstrated.

1. Kim WR, Brown RS Jr, Terrault NA *et al.* Burden of liver disease in the United States: summary of a workshop. *Hepatology* 2002; **36**: 227–242.
2. Ginés P, Quintero E, Arroyo V *et al.* Compensated cirrhosis: natural history and prognostic factors. *Hepatology* 1987; **7**: 122–128.
3. Marchesini G, Bianchi G, Amodio P *et al.* Italian Study Group for quality of life in cirrhosis. Factors associated with poor health-related quality of life of patients with cirrhosis. *Gastroenterology* 2001; **120**: 170–178.
4. Les I, Doval E, Flavià M *et al.* Quality of life in cirrhosis is related to potentially treatable factors. *Eur J Gastroenterol Hepatol* 2010; **22**: 221–227.
5. Moore KP, Aithal GP. Guidelines on the management of ascites in cirrhosis. *Gut* 2006; **55**: vi1–vi12.
6. Runyon BA. AASLD. Introduction to the revised American Association for the Study of Liver Diseases Practice Guideline management of adult patients with ascites due to cirrhosis 2012. *Hepatology* 2013; **57**: 1651–1653.
7. Gentilini P, Casini-Raggi V, Di Fiore G *et al.* Albumin improves the response to diuretics in patients with cirrhosis and ascites: results of a randomized, controlled trial. *Journal of Hepatology* 1999; **30**: 639–645.
8. Romanelli RG, La Villa G, Barletta G *et al.* Long-term albumin infusion improves survival in patients with cirrhosis and ascites: an unblinded randomized trial. *World J Gastroenterol* 2006; **12**: 1403–1407.
9. Okita K, Sakaida I, Okada M *et al.* A multicenter, open-label, dose-ranging study to exploratively evaluate the efficacy, safety, and dose-response of tolvaptan in patients with decompensated liver cirrhosis. *J Gastroenterol* 2010; **45**: 979–987.
10. Sakaida I, Yanase M, Kobayashi Y *et al.* The pharmacokinetics and pharmacodynamics of tolvaptan in patients with liver cirrhosis with insufficient response to conventional diuretics: a multicentre, double-blind, parallel-group, phase III study. *J Int Med Res* 2012; **40**: 2381–2393.
11. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; **30**: 473–483.
12. Bergner M, Bobbitt RA, Carter WB *et al.* The Sickness Impact Profile: development and final revision of a health status measure. *Med Care* 1981; **19**: 787–805.
13. Bayliss MS, Gandek B, Bungay KM *et al.* A questionnaire to assess the generic and disease-specific health outcomes of patients with chronic hepatitis C. *Qual Life Res* 1998; **7**: 39–55.
14. Younossi ZM, Guyatt G, Kiwi M *et al.* Development of a disease specific questionnaire to measure health related quality of life in patients with chronic liver disease. *Gut* 1999; **45**: 295–300.
15. Gralnek IM, Hays RD, Kilbourne A *et al.* Development and evaluation of the Liver Disease Quality of Life instrument in persons with advanced, chronic liver disease–the LDQOL 1.0. *Am J Gastroenterol* 2000; **95**: 3552–3565.
16. Unal G, de Boer JB, Borsboom GJ *et al.* A psychometric comparison of health-related quality of life measures in chronic liver disease. *J Clin Epidemiol* 2001; **54**: 587–596.

17. Fukuhara S, Bito S, Green J *et al.* Translation, adaptation, and validation of the SF-36 Health Survey for use in Japan. *J Clin Epidemiol* 1998; **51**: 1037–1044.
18. Embreston SE, Reise SP. *Item Respose Theory for Psycologists*. Lawrence Erlbaum Associates: Mahwah, NJ, 2000.
19. Muraki E. A generalized partial credit model: application of an em algorithm. *Applied Psychological Measurement* 1992; **16**: 159–176.
20. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. *Med Care* 1990; **28**: 632–642.
21. Garratt AM, Ruta DA, Abdalla MI *et al.* SF 36 health survey questionnaire: II. Responsiveness to changes in health status in four common clinical conditions. *Qual Health Care* 1994; **3**: 186–192.
22. Beaton DE, Hogg-Johnson S, Bombardier C. Evaluating changes in health status: reliability and responsiveness of five generic health status measures in workers with musculoskeletal disorders. *J Clin Epidemiol* 1997; **50**: 79–93.
23. Deyo RA, Centor RM. Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. *J Chronic Dis* 1986; **39**: 897–906.
24. van der Plas SM, Hansen BE, de Boer JB *et al.* Generic and disease-specific health related quality of life in non-cirrhotic, cirrhotic and transplanted liver patients: a cross-sectional study. *BMC Gastroenterol* 2003; **3**: 33.

Supplementary Information accompanies this paper on the Clinical and Translational Gastroenterology website (http://www.nature.com/ctg)